Submission to the Meeting of Joint Committee on Communications, Climate Action and Environment

Re: Online Advertising and Social Media (Transparency) Bill 2017

April 17th, 2018

Dr Ciarán Mc Mahon

Introduction

I would like to begin by pointing out two previous communications I have had with this House in relation to matters of politics and social media. In January 2013, in the wake of the death of Minister Shane McEntee, I sent a briefing paper to all members of the Dáil and Seanad. I'd like to quote from that paper, entitled *Politics, cyberbullying and social media: 6 psychological features of harmful electronic communication*' as I feel it remains very pertinent to today's conversation:

> However, the perception remains that cyberspace is a dark and awful place and this is not without foundation. The reality of this situation is at least partially as a consequence of a lack of enforcement of existing law in cyberspace

I concluded the paper as follows:

> Again, I call on elected representatives to lead by example - not only by using electronic communication technology in thoughtful and magnanimous ways – but also by ensuring that sufficient resources are made available so that Ireland's pre-eminent status as a European technology hub is paired with expertise in the cyberpsychological phenomena I have outlined above.

I returned to this them in a submission I made to the National Draft Risk Assessment exercise in 2014. In this paper I repeated this caution:

> … while in 2006 the New York Times labelled the IFSC the 'Wild West of European Finance', our capital is now host to a different type of risk entirely. Can we be sure that our tech giants are not engaging in practices that do not expose us to a similar level of risk? Critically, in the light of the 2007 cyberattacks on Estonia one would have to ask how exposed Ireland is, particularly with regard to how many tech companies have their European headquarters located here.

Obviously, I was not alone in making such warnings[1]. But here we are in 2018, witnessing the consequences of several years of light-touch regulation of social media. I note that an opinion pieces in the *Irish Times* argued that 'Ireland has failed to regulate Facebook on behalf of Europe'. Hence, at the outset, I must I commend Deputy Lawless in bringing this Bill to House – something of this kind is long overdue.

General comments

To paraphrase that old Chinese proverb, the best time to regulate the internet is 30 years ago, the second-best time is now. Essentially what has happened over recent decades is that in the absence of state-run services, multinational corporations have effectively created and monopolised online public spaces. However, it would be facile to 'blame' the social media corporations for the current problems. They have simply exploited the context that the state has allowed them to occupy and failed to regulate. Moreover, it is only by government and the social media corporations working together intensely can either thrive, let alone retain the confidence of the public. I stress that there are no technological solutions to psychosocial problems. It is only by a multi-factor strategy, using interdisciplinary expertise, across a number of sectors, will these risks be adequately mitigated.

Cyberpsychological aspects of social media context

Let me briefly review what I believe are the most pertinent observations on the cyberpsychological aspects of this context. *(See attached slides)*

Firstly, there are many contradictions in how people use online platforms. Here we the 'iceberg model'[2] of a person's information on an online platform. When a controversy erupts, users may change their privacy settings. But those settings only affect the visible part of the iceberg. However, the more valuable date is the much bigger part, under water, which they have little control over – except to perhaps delete their account. This is what has been termed the privacy paradox: while we might not like, for example, accepting a friend request from our boss on Facebook, we will continue to give the service much more valuable information by using it daily.

Secondly, I would like to give the Committee an idea of the scale of these problems. In this instance we have a visualisation of a network of Twitter bots, taken from a study[3] where the researchers examined the reported locations of a sample of recently created accounts. Ordinarily, one would expect this data to overlap with centres of population. But you can see from the rectangular shapes over North America and Europe that something unusual is taking place. These are obviously automatically created accounts. The researchers estimate that they represent a bot net of over 350,000 Twitter accounts. And this is only one example of a bot net, and one where the creators got sloppy with their coding. Its purpose is unknown, though probably commercial rather than political.

As an aside, recent research[4] estimates that 9 to 15% of all Twitter accounts are bots, meaning somewhere between 30 and 50 million accounts. In an earnings call on November 1, 2017[5], Facebook disclosed that of its 2.07 billion monthly active users, it estimated 10% of those were duplicate accounts and 2-3% were inauthentic or fake accounts. That's in the region of 124 to 207 million duplicate accounts, and 41 to 62 million fake accounts.

Thirdly, one often hears about content on social media 'going viral'. The idea here is that a message spreads like a virus, where once a person receives it, they pass it on to everyone they come into contact with, and so it cascades repeatedly, until it runs out of people to 'infect'. In fact, this is an illusion. Content on social media spreads more often in a traditional broadcast model, whereby large, central accounts – such as news organisations or celebrities 'infect' many people at once, who subsequently may or may not pass it on the message. Here we see some visualisations of large message cascades on Twitter[6] – as you can see, most of the propagation comes from one or two accounts, after which it terminates. Hence, we can't expect that the right information good news will simply 'go viral' once introduced to the network: it needs to be pushed from by someone with plenty of influence.

Fourthly, even when misinformation is corrected on social media, that doesn't necessarily stop it being propagated. Here is a graph from a study[7] of what happens when someone comments on a rumour on Facebook with a link to the Snopes debunking site. As you can see, there is very little difference between the ongoing re-sharing of rumours after they have been either confirmed as true or debunked as false by a Snopes link. While there is a drop-off, both continue to be shared to similar degrees. So, there is not much evidence for misinformation being naturally corrected, or even fading away on online social networks.

Fifthly, we come to the spreading of disinformation by adversaries. This is a graph of from a study[8] of the spread of a particular rumour in the wake of the 2013 Boston Marathon bombing. In this case the rumour was that the bomb was actually carried out by US Navy Seals as some kind of a 'false flag' operation. As you can see, while corrective reports circulated on Twitter very soon, the misinformation was still continued to be propagated. By all accounts, it should have dropped in intensity, as the previous examples, but as you can see it actually spikes.

The researchers in question suspect that this is evidence of adversaries deliberately pushing a disruptive message.

Finally, although I don't have a graph to illustrate this – it has been demonstrated empirically[9] that being able to target users most psychologically vulnerable to disinformation is key to its propagation across social media. Moreover, while there is obviously no extant research on the current situation in Ireland, it would be naïve to think that this is not occurring, nor that it will not occur in the future. The enemies of the open society simply wish to spread confusion and division – they are not interested in any ideology, other than the disruption of democracy. The fundamental point for legislators to realise from Slides 1, 2, 3, 4, and 5 is that unregulated online platforms have allowed the creation of a cyberpsychological environment in which it is increasingly difficult for voters to know what is true and who can be trusted. Consequently, it behoves both Government and the social media corporations to work together, as quickly as possible, to rectify this situation.

Moreover, let us dispense with the idea that social media has no effect on election. We already know, by Facebook's own account, that it does. A 2012 paper published in *Nature* using Facebook data on 61 million American users[10], showed that Facebook messaging influenced actual voter behaviour. In this case, when people saw that their friends using a simple 'I voted' badge, they themselves were more likely to vote. Hence, as has been argued before, Facebook could very easily swing an election by showing these badges in some constituencies and not others. This case was well argued by Harvard law professor Jonathan Zittrain[11] long before the 2016 election, though few were paying attention at the time.

Specific points regarding the Bill

Part 1, Section 2

Firstly, I would like to query Part 1, Section 2(2), in relation to the Referendum Commission. While I gather that this is an exercise in definitions, I would urge the committee to consider that State communication in social media as a necessary component in dealing with misinformation. As I have stated above, the nature of social media means that disinformation not only will continue to be propagated, as well as manipulated by adversaries. Consequently, in combating disinformation campaigns, it is essential that the State is prepared to engage in ongoing factual education on social media. I will return to this point in the Concluding remarks below.

Part 2, Section 3

Secondly, in relation to Part 2, Section 3(2), again I would ask the Committee to give serious consideration to the State and its own information campaigns, more of which I will discuss below. Furthermore, I would like to remind the Committee that, none of this will apply to 'organic' or unpaid content, which, as I have shown above, is far more common than advertising. As politicians you will know that arguably most politically damaging activity in the 2016 US Presidential election was not social media advertising, but hacking. The exfiltration and selective leaking of the emails of the Democratic National Committee would probably not have happened had its members practiced basic cybersecurity, specifically two-factor authentication. In that light, I would advise members to consult the Cybersecurity Campaign Playbook[12] recently published by the Belfer Center for Science and International Affairs, at the Harvard Kennedy School.

Part 2, Section 4

Thirdly, with regard to Section 4 of this bill, and its definitions of the transparency notice, a few points can be made. In relation to Section 4(1)(b), and the elements of the transparency notice, I would suggest that it contain not only whatever website that these ads are linked to, but that the online platforms are required to ensure that these are not click-through websites – i.e. deceptively linked to a first website, but subsequently redirected to second and possibly third website. These kinds of deceptions are common in disinformation campaigns.

It is also essential that the online platforms clarify when demographic, socio-economic, ethnic and/or psychographic targeting is used in these advertisements. In this regard, I would draw the Committee's attention to two reports from the independent investigative newsroom, ProPublica. In October 2016, its reporters revealed that it was possible to buy ads on Facebook which were targeted by ethnicity[13]. In other words, landlords could advertise to exclude potential black or Hispanic tenants. Naturally, this created quite a controversy, and in a press release dated November 11, 2016, Facebook stated that it took the issue of ethnic discrimination very seriously, and that it was introducing an automated detection system to prevent this from happening again. However, on November 21, 2017, ProPublica published another report[14], which showed that its reporters were still able to place similar discriminatory advertisements.

As such, I would urge the Committee to seek significant detail from the online platforms regarding their ad targeting capabilities, and how those capabilities comply with Irish and European anti-discrimination laws. Moreover, as their ad targeting capabilities develop, the online platforms should be obliged to register those changes with statutory authorities on an ongoing basis.

With regard Section 4(3), and the transparency notices, my estimation is that, in practice, it will be quite difficult in many cases to make online political ads visibly transparent (i.e. there may simply not be enough space for legible text). As such, I suggest that not only should the online platforms maintain central registers of all the political ads they carry, but that the State create its own independent cross-platform authority where ads must be registered also. This could also register newspaper ads and political broadcasts. I note that the establishment of an Electoral Commission is listed as a matter of priority in the 2016 'A Programme for a Partnership Government'[15]. Such a Commission would be an excellent fit for this registry.

Part 2, Section 5

Fourthly, in relation to Section 5, I draw the Committee's attention to an item currently before Congress in the United States, S.1989[16]. The Honest Ads Act is similar to the current Bill, in that it requires more transparency in online political advertising. Moreover, in the light of recent events, despite some mixed signals, it does now appear that both Twitter and Facebook have stated publicly that they support this Act. However, it should be noted that the Honest Ads Act also contains provisions for penalties for the online platforms. This Bill does not have such provisions, which I believe is an error. There should be a financial penalty for an online platform which carries political ads with no transparency notices, and that penalty should be proportionate to the platform's publicly claimed monthly active users.

It is worth noting some announcements made by social media corporations since last autumn. At the outset, much of what is in these proposed policy changes is similar to what is proposed in this Bill – more transparency, traceability and authenticity and so on. Hence, it would be very surprising to me if they were to oppose this Bill in any meaningful way.

In terms of those companies' individual announcements, some further detail is useful. On October 24th of last year, Twitter announced that it would 'within weeks'[17] launch 'an industry-leading transparency center that will offer everyone visibility into who is advertising on Twitter, details behind those ads'. There are many elements of this announcement which

I'm sure we would all welcome, including transparency about the identity of advertisers, their total spend, and who they target, as well as stricter limits on electioneering advertising.

However, this 'Advertising Transparency Center' has not actually been launched in the intervening six months since it was announced, and last week Twitter said it was committing to launching it 'this summer'[18]. Ireland is a small country, which has been very good to Twitter, so I would suggest that, as demonstration of its commitment to her democracy, Twitter prioritise the launch of this Advertising Transparency Center here, by the end of this month.

Furthermore, on October 27 last year[19], and April 6 this year[20], Facebook announced many changes as to how it would manage transparency in political advertising. Again, these changes would be most welcome – being able to 'View Ads' for every page, and political advertisers having to verify their identity, as well as detailed transparency on every ad.

However, despite saying in October that these features would be 'starting next month' after a test in Canada, this feature is still not available, except in Canada. And in April, Facebook said its plan was now to launch it globally in June. Once again, I repeat that Ireland is a small country, which has been very good to Facebook, so I would suggest that, as demonstration of its commitment to both this country and her democracy, Facebook prioritise the launch of this View Ads and related features here, by the end of this month.

Another detail is worth pointing out here. I believe that in revising this Bill that the Committee should insist that the oversight of online political advertising be handled locally, and by actual humans. That is, in its statement of October 27, 2017, Facebook stated, "For political advertisers that do not proactively disclose themselves, we are building machine learning tools that will help us find them and require them to verify their identity." This is also emphasised in its statement of April 6th, with reference to 'investment in artificial intelligence' to check ads.

I believe that this is an inappropriate supervision method as machine learning tools have shown themselves to be inherently biased[21]. Additionally, as the second ProPublica report found, Facebook's automated detection of discriminatory ads does not appear to have worked. Again, I repeat: technical solutions to psychosocial problems are not viable in the long-term. I would urge the Committee to insist that all political ads being served in Ireland by online platforms be overseen by locally employed specialists. Moreover, these should be full-time employees in Ireland – not poorly paid contractors on the far side of the world as is too often the case in content moderation[22]

This country is not so big that this cannot be easily achieved. If either Facebook or Twitter or any other online platform cannot afford to employ enough people to manually oversee its political ads here, then it should not be accepting cash from those accounts in the first place.

Part 2, Section 6

Fourthly, in relation to Section 6, and the offence of using a bot – obviously, this assumes that that creator of the bot network can be traced, and indeed charged. As I have outlined above, from the public side of social media, that strikes me as Herculean task. I would advise the Committee to take a look at the Computational Propaganda Project[23] recently concluded by the Oxford Internet Institute. This interdisciplinary study is a fascinating exploration of the variety of political bot activity across the globe, and very revealing of what can happen online in different political contexts. The study covered 9 countries, although Ireland was not one of them.

In personal correspondence with Professor Phillip Howard, one of the co-authors of that report, he agrees that policy guidance setting that a bot must identify itself as such would be welcome, but that a full bot ban would probably be neither healthy or viable.

As such, I draw the Committee's attention to an item currently before the California legislature, Bill AB-1950[24] (Internet Web sites: social media: advertising: accounts). This bill puts the onus on the online platforms to identify and label automated or bot accounts. I think this would be a better approach, and more productive in the long term.

Concluding remarks

Ultimately, while this bill contains many elements that will certainly prove useful in delivering transparency in political social media, it should form part of a broad, multidimensional and interdisciplinary approach. As such, it should really be seen as the first wave in a rolling project to regulate online platforms for the betterment of democracy. It would not be beyond the realms of possibility that these measures might need to be revisited, with further actions. I note that in other arenas more punitive measures have been suggested, including banning targeted political advertising outright, enforcing opt-in for data usage and indeed breaking up social media monopolies. I remind the Committee that the European Commission fined Facebook €110 million in May 2017[25] for providing misleading information during its purchase of WhatsApp. Legislators must be more vigilant, and service providers must be more transparent.

Moreover, I believe that, in addition to amending and progressing this Bill, the State should prioritise two additional projects. Firstly, let me quote from the European Commission's recent Report of the Independent High-Level Group on fake news and online disinformation[26]. The following actions from its conclusions are worth noting:

- promote and sharpen the use of media and information literacy approaches to counter disinformation and help users navigate our digital information environment,
- develop tools for empowering users and journalists and foster a positive engagement with fast-evolving information technologies

- calibrate the effectiveness of the responses through continuous research on the impact of disinformation …

Similarly, a joint letter in *Science* published last month by a broad group of experts, led by Professor David Lazer of Northeastern University, echoes these sentiments.:

> Our call is to promote interdisciplinary research to reduce the spread of fake news and to address the underlying pathologies it has revealed… More broadly, we must answer a fundamental question: How can we create a news ecosystem and culture that values and promotes truth?

Hence, I would urge the Committee to consider funding a multidisciplinary basic research project as per the recommendation of both of these publications. There are plenty of such research programmes underway in other countries, but none here. This is an acute deficiency given how many multinationals have their EMEA headquarters here, who I am sure would be eager to contribute to such a project.

Secondly, as alluded to above, I note the Bill stipulates that no public monies be spent on political advertising. Yet defines political advertising in a quasi-negative sense - as in, whatever the Referendum Commission does, it is not engaging in political advertising. While I do not wish to re-litigate the Strategic Communications Unit controversy, we do need to be think again how the State communicates to its citizens - whether it be through the Referendum Commission, Citizens Information, Merrion Street.ie, or RTÉ.

In fact, I suggest that in the light of ongoing global disinformation campaigns, the State must significantly increase its educational, information campaigns during referendum and election campaigns. I strongly advise the Committee to recognise that, even if this Bill, including the suggestions I have made above, were to succeed, that would be no guarantee of an optimal informational context occurring – keeping bad information out, does not necessarily imply that factual information will magically appear. As I have said, the Bill will not apply to organic social media content. And again, I stress that politicians and their teams need to improve their own operational cybersecurity practices.

Hence, the State needs to be prepared to mount an ongoing factual information operation via the online platforms. This will require social media advertising, though this should not require an extraordinary sum of money. For example, Facebook disclosed in September 2017 that approximately a total of $100,000 had been spent on ads by Russian trolls[27]. Similarly, it was reported that Twitter offered 15% of all its US election advertisements to the Russia Today organisation for $3 million[28]. To compare, in its most recent report, the Referendum Commission reports that it spent €132,000 on online advertising and €56,000 on 'new media expenditure' over the course of the last two referendums[29].

As such, an annual spend of in the region of €1 million should suffice to combat whatever misinformation campaigns the State would encounter – and again, the online platforms should, as part of their ongoing commitment to democracy, give the State a substantial discount here. Again, the promised Electoral Commission would be an excellent fit for these rolling campaigns.

Finally, I thank the Committee for the invitation and wish it every success in dealing with these most pressing of issues for our democracy. I call on the online platforms to engage with the suggestions above in a transparent and mature spirit. It is only by working together that online environments can be restored to a level of truthfulness and trustworthiness that our citizens deserve, and our children will inherit.

Notes

[1] https://www.irishtimes.com/news/ireland/irish-news/facebook-privacy-flaw-was-flagged-with-irish-regulator-in-2011-1.3432013

[2] See Slide 2 of the attached presentation or Debatin, B., Lovejoy, J. P., Horn, A.-K., & Hughes, B. N. (2009). Facebook and online privacy: Attitudes, behaviors, and unintended consequences. Journal of Computer-Mediated Communication, 15(1), 83–108. https://doi.org/10.1111/j.1083-6101.2009.01494.x

[3] See Slide 3 of the attached presentation or Echeverría, J., & Zhou, S. (2017). Discovery, Retrieval, and Analysis of "Star Wars" botnet in Twitter. https://doi.org/10.1145/3110025.3110074

[4] See Varol, O., Ferrara, E., Davis, C. A., Menczer, F., & Flammini, A. (2017). Online Human-Bot Interactions: Detection, Estimation, and Characterization. http://arxiv.org/abs/1703.03107

[5] See 3Q Transcript, https://investor.fb.com/financials/default.aspx , here https://s21.q4cdn.com/399680738/files/doc_financials/2017/Q3/Q3-'17-Earnings-call-transcript.pdf

[6] See Slide 4 of the attached presentation, or Goel, S., Watts, D. J., & Goldstein, D. G. (2012). The structure of online diffusion networks. In Proceedings of the 13th ACM Conference on Electronic Commerce - EC '12 (p. 623). New York, New York, USA: ACM Press. https://doi.org/10.1145/2229012.2229058

[7] See Slide 5 of the attached presentation, or Friggeri, A., Eckles, D., & Cheng, J. (2014). Rumor Cascades. In Eighth International AAAI Conference on Weblogs and Social Media. Ann Arbor, MI: AAAI Publications.

[8] See Slide 6 of the attached presentation, or Starbird, K., Maddock, J., Orand, M., Achterman, P., & Mason, R. M. (2014). Rumors, False Flags, and Digital Vigilantes: Misinformation on Twitter after the 2013 Boston Marathon Bombing. IConference 2014 Proceedings. https://doi.org/10.9776/14308

[9] Aymanns, C., Foerster, J., & Georg, C.-P. (2017). Fake News in Social Networks. https://arxiv.org/pdf/1708.06233

[10] https://www.nature.com/articles/nature11421

[11] See https://newrepublic.com/article/117878/information-fiduciary-solution-facebook-digital-gerrymandering

[12] https://www.belfercenter.org/CyberPlaybook

[13] https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race

[14] https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin

[15] https://www.merrionstreet.ie/MerrionStreet/en/ImageLibrary/Programme_for_Partnership_Government.pdf

[16] https://www.congress.gov/bill/115th-congress/senate-bill/1989/text

[17] https://blog.twitter.com/official/en_us/topics/product/2017/New-Transparency-For-Ads-on-Twitter.html

[18] https://twitter.com/Policy/status/983734917015199744

[19] https://newsroom.fb.com/news/2017/10/update-on-our-advertising-transparency-and-authenticity-efforts/

[20] https://newsroom.fb.com/news/2018/04/transparent-ads-and-pages/

[21] See keynote address by Kate Crawford https://www.youtube.com/watch?v=fMym_BKWQzk

[22] See Adrian Chen in *Wired* https://www.wired.com/2014/10/content-moderation/. Also short documentary, *The Moderators*, https://vimeo.com/213152344

[23] http://comprop.oii.ox.ac.uk/research/working-papers/computational-propaganda-worldwide-executive-summary/

[24] https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB1950

[25] http://europa.eu/rapid/press-release_IP-17-1369_en.htm

[26] https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation

[27] https://www.nytimes.com/2017/09/06/technology/facebook-russian-political-ads.html

[28] https://www.buzzfeed.com/alexkantrowitz/twitter-offered-rt-15-of-its-total-share-of-us-elections

[29] http://www.refcom.ie/previous-referendums/marriage-presidential-age/report-on-the-referendums-on-marriage-and-on-the-age-of-presidential-candidates.pdf